Sponsored by:

Professional Association for SQL Server

24 HOURS OF PASS

FIVE Physical Data Modeling Blunders .... and How to Avoid Them

Karen Lopez

Speaker Bio

Karen López is a principal consultant. She specializes in the practical application of data management principles. Karen is also the ListMistress and moderator of her company’s Discussion Groups at www.infoadvisors.com and dm-discuss (Yahoo Groups).
Agenda

- The Problem
- Blunders, FAILs, D’Oh’s, Faults and WTHs?
- 10 Tips

POLL: Who Are You?
POLL: Physical Model Much?

The Problem
Assuming Numbers are Numbers

D’OH!

Leading Zeros
Non-Numeric
Special Characters
Sorting Issues
Externally Managed Numbers
Avoid it

1. Use a numeric-ish datatype when there’s math involved.
2. Do your research
3. If it’s externally managed data, its format or composition could change at any time. Anticipate that.
4. Anticipate leading zeros.
5. Anticipate significant formatting, characters & other tricks.
Choosing the Wrong Primary Key

32 + 8 + 32 + 32 + (0 to 22) = A really big key that only a Data Architect could love
Avoid it

1. Establish Primary Key standards and styles
2. Smaller PKs lead to smaller indexes and smaller databases. Take advantage of that
3. Choose PKs that do not change when data changes
4. Understand how Identity columns work

Applying Surrogate Keys Incorrectly
Applying a Surrogate Key

WTH?
Alternate Key

1. Establish Primary Key standards and styles
2. Smaller PKs lead to smaller indexes and smaller databases. Take advantage of that.
3. Don’t forget the Business Key
4. Choose PKs that do not change when data changes
5. Understand how Identity columns work

Avoid it
Turning off RI for Development

Blunder
RI = FK

Avoid it

Educate team on the risks & nearly universal outcome of turning off RI
Generate RI in every script

Develop test data
Develop scripts for loading data
Ensure team has easy access to Data Models
Using the Defaults

Defaults <> Best Practices
Product Defaults aren’t YOUR defaults
Defaults might even be...faulty.
Avoid it

1. Create a Generation Test database...and use it.
2. Test Generation Options to Test DB with Data
3. Generate Test Scripts, Varying Options
4. Review options with DBAs and Developers.
5. Choose Defaults, don’t default them
6. Save your Default Set
Other Blunders

- Using the wrong tool
- Too much Subtyping
- Overly Flexible hierarchies
- Wrong Datatypes
- Forgetting to manage the data
- Duplicate / Unusable Indexes
- Ignoring Clustering
- Using overly generic design patterns
- Column order tango
- Silly Long Names
- Failing to compare
- Hand generating scripts
- Failing to report issues
- Denormalizing too soon
- Failing to use constraints
- Making performance more important than integrity
- Completely separate logical model
- Failing to take into account whole environment
- Applying surrogate keys where they have no business
- Not testing your design
- Ignoring reference data design

10 Tips for Avoiding Physical Modeling Blunders

1. Understand cost, benefit and risk
2. Get formal training on target technologies
3. Be physically present where your designs are used.
4. Ask about trade-offs, not just solutions
5. Build portfolio of performance vs. integrity trade-offs
10 Tips for Avoiding Physical Modeling Blunders

6. Profile Source Data
7. Build Test Databases, with Data
8. Test Scripts
9. Compare, Compare, Compare…then Compare some More.

Summary

1. Don’t assume that numbers are….numbers
2. Choose the right primary key
3. Apply surrogate keys correctly
4. Keep RI turned on
5. Architect, don’t default
Questions?

Feedback, questions, comments are always appreciated.

http://www.speakerrate.com/karenlopez

Karen@InfoAdvisors.com

@datachick
facebook.com/lopezk
linkedin.com/in/karenlopez

Professional Association for SQL Server

Thank you to our sponsor

Dell
May 11-13, Orlando, FL

Save 25%: Register by April 12th
www.sqlpass.org/sqlrally

Oct 11-14, Seattle, WA

Register by March 31st: save 40% and have the chance to win a cruise to Alaska!
www.sqlpass.org/summit